THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 15, Issue of April 13, pp. 11913–11921, 2001
Printed in U.S.A.

The *Caenorhabditis elegans unc-32* Gene Encodes Alternative Forms of a Vacuolar ATPase *a* Subunit*

Received for publication, October 17, 2000, and in revised form, November 30, 2000
Published, JBC Papers in Press, December 7, 2000, DOI 10.1074/jbc.M009451200

Nathalie Pujol‡§¶, Claire Bonnerot§, Jonathan J. Ewbank‖, Yuji Kohara, and Danielle Thierry-Mieg§‡‡**

*From the ‡Laboratoire de Génétique et Physiologie du Développement, CNRS, INSERM, Université de la Méditerranée, Luminy Case 907, 13288 Marseille Cedex 9, France, §Centre de Recherches en Biochimie Macromoléculaire, CNRS, 34293 Montpellier Cedex 5, France, ‖Centre d'Immunologie de Marseille-Luminy, CNRS, INSERM, Université de la Méditerranée, 13288 Marseille Cedex 9, France, and **National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan*

Eukaryotes possess multiple isoforms of the *a* subunit of the V_0 complex of vacuolar-type H^+-ATPases (V-ATPases). Mutations in the V-ATPase *a3* isoform have recently been shown to result in osteopetrosis, a fatal disease in humans, but no function has yet been ascribed to other isoforms. In *Caenorhabditis elegans*, the *unc-32* mutant was originally isolated on the basis of its movement defect. We have isolated four new mutant alleles, the strongest of which is embryonic lethal. We show here that *unc-32* corresponds to one of the four genes encoding a V-ATPase *a* subunit in the nematode, and we present their expression patterns and a molecular analysis of the gene family. *unc-32* gives rise via alternative splicing to at least six transcripts. In the uncoordinated alleles, the transcript *unc-32* B is affected, suggesting that it encodes an isoform that is targeted to synaptic vesicles of cholinergic neurons, where it would control neurotransmitter uptake or release. Other isoforms expressed widely during embryogenesis are mutated in the lethal alleles and would be involved in other acidic organelles. Our results indicate that V-ATPase *a* subunit genes are highly regulated and have tissue-specific function.

Vacuolar-type H^+-ATPases (V-ATPases[1]) are ATP-dependent proton pumps involved in the acidification of intracellular compartments in eukaryotic cells and are important for many facets of cellular and organismal function (for review, see Refs. 1–3). As well as contributing to fundamental cellular processes in the majority of eukaryotic cells, they also play specialized functions in certain cell types. In macrophages, the progressive acidification of phagosomes is an integral part of their maturation into lysosomes, whereas the acidification of early endocytic compartments is necessary for the recycling to the plasma membrane of transport molecules such as the transferrin receptor. In neurons, neurotransmitter trafficking is dependent on V-ATPases. The H^+-driven antiporters responsible for synaptic vesicle import of a specific neurotransmitter, such as the vesicular acetycholine transporter, are known to be driven by proton gradients established by a V-ATPase localized to the same vesicle (4, 5).

The V-ATPases are composed of two complexes, V_1 and V_0, that serve different functions (for review, see Refs. 6, 7). V_1, a complex of 570 kDa composed of eight subunits (*A–H*), is located on the cytoplasmic side of the vesicle membrane and is responsible for ATP hydrolysis. V_0 is a 260-kDa complex with five different subunits (*a*, *d*, *c*, *c′*, and *c″*) of 17–100 kDa. It is an integral membrane complex and is responsible for proton translocation from the cytoplasm to the lumen of the vesicle. In essence, the V-ATPases resemble, both in terms of sequence and structure, the ATP synthases of mitochondria. ATP synthases, however, couple the synthesis of ATP to the flow of protons from the mitochondrial matrix to the cytoplasm, and the overall topology of its two components is inverted, with the F_1 complex (analogous to the V_1 complex) being on the lumenal side. Interestingly, there is no clear homologue for the V_0 100-kDa *a* subunit within the analogous F_0 complex, suggesting that it plays a specific role in the function of the V-ATPases. In yeast, there are two *a* subunit genes that encode proteins that have distinct subcellular localization (8). One, VPH1p, is associated with the central vacuole, whereas STV1p is thought to be a Golgi/endosome-specific form. Thus the two *a* subunits may be involved in targeting yeast V-ATPases to different intracellular membranes. Three murine *a* subunit genes have been characterized (*a1*, *a2*, and *a3*; Ref. 9). For *a1* and *a3*, there are two and three alternative transcripts, respectively, giving the possibility of substantial variation in the structure of the murine V-ATPase. Intriguingly, the level of expression of the different transcripts of the murine genes varies among different tissues (9). This suggests that the V-ATPase complexes may have specialized compositions and structures, not only at the intracellular level but also in different tissues. The recent characterization of knockout mice for the *a3* gene (ATP6i), which exhibit severe osteopetrosis due to loss of osteoclast-mediated extracellular acidification but no general defect in V-ATPase function, supports such a hypothesis (10). More recently, mutations in the human orthologue have been shown to

* This work was supported by institutional grants from Ministère de la Recherche et de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale, and Centre National de la Recherche Scientifique. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF320899, AF320900, AF320901, AF320902, AF320903, and AF320904.

¶ To whom correspondence should be addressed: Laboratoire de Génétique et Physiologie du Développement, Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, Université de la Méditerranée, Luminy Case 907, 13288 Marseille Cedex 9, France. Tel.: 33-49126-9733; Fax: 33-49126-9726; E-mail: pujol@ibdm.univ-mrs.fr.

‡‡ Present address: NCBI, NLM, National Institutes of Health, Bethesda, MD 20894.

[1] The abbreviations used are: V-ATPase, vacuolar-type H^+-ATPase; bp, base pair; PCR, polymerase chain reaction; RT-PCR, reverse transcription-polymerase chain reaction; SL, spliced leader; GFP, green fluorescent protein; Unc, uncoordinated.

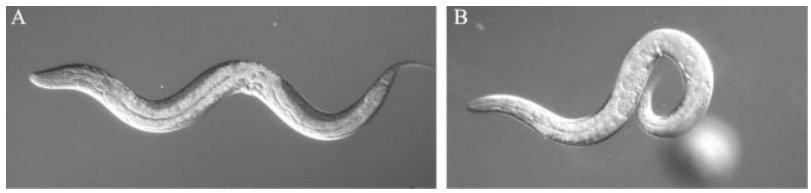

FIG. 1. **The *unc-32* uncoordinated mutant phenotype.** Differential interference contrast images of wild-type (*A*) and *unc-32(e189)* mutant (*B*) worms as young adults are shown. *unc-32* mutants are characterized by a reverse ventral coiler phenotype.

be responsible for a subset of human osteopetrosis (11, 12).

Here, we describe the genetic and molecular identification of the *Caenorhabditis elegans unc-32* gene, which encodes a set of V-ATPase 100-kDa *a* subunits. We show that the gene is one of a family of four V-ATPase *a* subunits encoding genes expressed in *C. elegans*. Furthermore, we show that a number of alternative *unc-32* transcripts are produced by differential splicing of alternative exons and have characterized five mutants, two of which affect specifically one isoform. On the basis of our phenotypic and molecular analysis of the mutants, we postulate that specific isoforms may be involved in nervous system function, whereas others are required for viability.

EXPERIMENTAL PROCEDURES

Strains and Mutagenesis—All strains including the wild-type N2 Bristol and strains containing the *unc-32* alleles and the transformants were grown and maintained at 20 °C as described (13). Four new ethyl methanesulfonate-induced alleles, *unc-32(f120), unc-32(f121), unc-32(f123),* and *unc-32(f131)*, were isolated in two noncomplementation screens for new *unc-32* alleles. The three former were induced on a *dpy-17(e164) ncl-1(e1865)* chromosome, the latter on an *unc-36(e251)* chromosome. Several thousand mutagenized genomes were tested against *unc-32(e189)* and scored for an uncoordinated phenotype. The new alleles were back-crossed at least four times to wild type, and genetic analysis was conducted using standard procedures. The two alleles, *f121* and *f123*, were homozygous lethal and were kept heterozygous with the balancer *qC1, dpy-19 glp-1*, provided by J. Kimble. The integrated transformant *eIs16*[*ZK637, unc-32*+] was kindly provided by J. Sulston. The nonsense suppressor *sup-7(st5)* (14) as well as the *smg-2(e2008)* suppressor involved in mRNA surveillance (15, 16) was combined with all alleles but *f131*.

Plasmid Constructions—All the plasmids used in this work were derived from the A1E5 plasmid (a gift from M. Craxton). A1E5 is a partial *Sau*IIIA restriction fragment of the ZK637 cosmid (GenBank accession number Z11115), from 19434 to 32570, cloned in BlueScribe. E5-NS was obtained after *Nhe*I and *Spe*I restriction and religation. E5-BK was constructed by a partial *Bam*HI restriction to keep the *Bam*HI site of the large exon of ZK637.8. E5ΔPac corresponds to a deletion between the two *Pac*I restriction sites of A1E5. E5-lin-9 was derived from E5-BK by *Sma*I and *Sac*II restriction, T4 DNA polymerase treatment, and religation. E5-BK*, introducing a frameshift in exon 6, was derived from E5-BK by *Bam*HI restriction, Klenow treatment, and religation. The green fluorescent protein (GFP) fusion in exon 4b, *unc-32*e4::GFP, was constructed by introduction of a 3600-bp *Sma*I-*Pst*I restriction fragment from E5-NS in the *Hin*DIII (blunted)-*Pst*I-digested plasmid pPD95.81 (a gift from A. Fire).

Transgenic Lines—Plasmids for injection were prepared by a standard alkaline lysis procedure, followed by LiCl precipitation. Injections were performed as described (17). Subclones of ZK637 were injected at a concentration of 10 ng/μl with plasmid pRF4 (encoding *rol6(su1006),* which confers a dominant rolling phenotype) at a concentration of 100 ng/μl. Some of the transformants also received either plasmid pPD93.97 (encoding a GFP reporter gene driven by the *myo-3* promoter) or pPD20.97 (encoding a LacZ reporter gene driven by the *myo-2* promoter) at a concentration of 20 ng/μl to check the presence of coinjected DNA directly by screening GFP in body wall muscle or β-galactosidase in the pharynx by 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside staining (both pPD plasmids were gifts from A. Fire). Clones were injected into *e189* or *f131* homozygotes and in *unc-32(e189* or *f123* or *f121) dpy-17(e164)/qC1*.

Transcript Analysis—Partial sequences from 15 cDNAs were aligned on the genome, and three clones were fully sequenced (yk15a11, yk244c11, and yk5d7). There was evidence for alternative splicing, and the alternative transcripts from the three clones were designated as *unc-32* A, B, and C, respectively. The differential splicing patterns were also analyzed by restriction digestion and/or sequencing of specific reverse transcription-polymerase chain reaction (RT-PCR) products. Total RNA from mixed populations of wild-type worms was prepared as previously described (18), reverse transcribed, and subjected to PCR with various sets of primers as well as the transpliced leaders SL1 and SL2 (19, 20). Amplification conditions and primers are available on request. The RT-PCR products were analyzed by restriction using enzymes cutting the transcripts only once, in exon 3 (*Xho*I), exon 4a (*Eco*RI), exon 4b (*Pst*I), exon 7a (*Pvu*I), and exon 7b (*Drd*I). Some RT-PCR products were sequenced on an ABI machine using dye deoxy-terminators (Prism; PerkinElmer Life Sciences). The traces were analyzed using the *acembly* program, developed by J. Thierry-Mieg and M. Potdevin.

When referring to the ZK637 cosmid, exon positions are: exon 1, 21653–21816 (coding sequence starting at 21664); exon 2, 21898–22074; exon 3, 22182–22310; alternative exon 4a, 22560–22714; alternative exon 4b, 23003–23109; alternative exon 4c, 23347–23468; exon 5, 23629–23838; exon 6, 23893–25147; alternative exon 7a, 25269–25391; alternative exon 7b, 26117–26257; exon 8, 26370–26589; exon 9, 26754–26925; and exon 10, 27145–27250. Transcripts were transpliced with SL1 at position 21653. Alternative polyadenylation occurs at similar frequencies at positions 27671 (atypical poly(A) signal agtaaa 16–11 bp before the poly(A)) or 27784 (atypical poly(A) signal aacaaa 17–12 bp before the poly(A)). Note that the existence, extent, and alternative status of exons 4a, 4b, 4c, 7a, and 7b were wrongly predicted for ZK637.8 in the data bases.

In Situ and GFP Expression Analyses—In situ hybridization analyses were performed as previously described (21). Transgenic worms carrying the *unc-32*e4::GFP construct were observed with a confocal microscope; the image in Fig. 6*G* is the projection of several *Z* serial sections.

Mutant Analysis—The genomic DNA of four mutant alleles (all but *f123*) was amplified with Takara *Taq*. PCR products covering the segments from −629 to 832, 394 to 1761, 1687 and 2162 to 3804, 2768 to 4622, and 4605 to 5851 (where position 1 is the A of the initiator ATG) were fully sequenced on both strands. RNA was extracted from the three alleles mutated in a splice junction, *e189, f120,* and *f131*, and subjected to RT-PCR analysis (see above) and sequencing.

Sequence Comparisons—The nematode VHA-5 and VHA-6 proteins correspond respectively to the predicted F35H10.4 (GenBank accession number AAA81682) and VW02B12.1 (GenBank accession number CAA90758) proteins, encoded by the expressed YK6569 and YK634 genes. The nematode VHA-7 protein, corresponding to the predicted C26H9A.1 protein (GenBank accession number T19492), encoded by the YK778 expressed gene, was edited by removing the first 291 amino acids and the next to last exon, which does not occur in 5 of 5 independent cDNA clones. The 5′ region removed does not correspond to any available cDNA, nor does it resemble sequence in any of the well characterized 100 kDa *a* subunit proteins. Comparisons between the available human genomic sequence (GenBank accession number AC002537) and the published cDNA sequence for the human *a1* gene (GenBank accession number NM_005177) allowed the deduction of the intron-exon boundaries. The intron-exon boundaries for the human *a3* gene (OC116) were published previously (22). Amino acid sequences were aligned with the Multiple Alignment Program (developed by X. Huang). Phylogenetic trees were produced using the full-length protein sequences by neighbor joining using Clustalw alignment in Phylip format and the Treeview program.

RESULTS

Three unc-32 Alleles Are Viable but Uncoordinated—The *unc-32(e189)* allele was isolated by Brenner (13) as part of the first genetic screen for *C. elegans* mutants. Homozygous mutants are healthy and of normal size but exhibit a ventral coiler

TABLE I

Complementation, dosage effects and molecular defects in unc-32 alleles

The phenotypes of all the combinations between the different *unc-32* alleles were analyzed. Let, lethal; Mel, maternal effect lethal; Emb, embryonic; Ste, sterile; WT, wild-type; ND, not determined. *f120* is molecularly and genetically identical to *e189*. All the alleles were combined with the deficiency of the entire locus *nDf17*, indicating that the null phenotype is embryonic lethal, the same as seen in *f123* homozygotes.

	unc-32 alleles			
	e189 or f120	*f131*	*f121*	*f123*
e189 or *f120*	Unc L2 to adult			
f131	Strong Unc	Extreme Unc, all stages		
f121	Almost WT	Strong Unc, especially in larvae	L2, L3, or L4 Let	
f123	10% extreme Unc, 90% L1 Let	L1 Let	L1 Let	Emb Let
nDf17	10% extreme Unc, 90% Let	L1 Let	L1 Let	Emb Let
unc-32+ Transformants[a]	WT	Unc+, some Mel, Ste	Let+, Unc+, but Mel, Ste	Let+, Unc+, but Mel, Ste
sup-7 interactions[b]	None	ND	None	Emb Let suppressed
Nature of allele	Hypomorph	Hypomorph	Hypomorph	Null
Molecular lesion[c]	g1339a Alternate exon 4b splice junction	g3485a Exon 6 splice junction	g2874a Gly[481/465/470] Glu	ND Stop likely

[a] Transgenic worms with constructs containing the wild-type *unc-32* locus. The Unc and Let phenotypes are rescued, but for the *f121* and *f123* alleles, as well as for *f131* to a lesser extent, Mel and/or Ste (abnormal gametogenesis) phenotypes are revealed.

[b] *sup-7* can suppress the embryonic lethality of *f123*, suggesting that this allele contains a premature stop codon.

[c] Positions of the molecular lesion, identified by sequencing the different alleles (position 1 is on the A of the initiator ATG); the position of the Gly amino acid mutated in the *f121* allele is given for the different proteins containing, respectively, exons 4a, 4b, and 4c.

phenotype during backward movement (Fig. 1). This defect is seen from the L2 stage onward. The only additional defect observed in *e189* is the brood size, smaller by one-third when compared with wild type, probably because of a reduction in number of fertile sperm. We carried out a genetic noncomplementation screen and isolated two more recessive viable *unc-32* alleles, *f120* and *f131.* Although *f120* shows a similar phenotype to *e189* and was subsequently found to correspond to an identical molecular lesion (see below), *f131* presents a much stronger uncoordinated (Unc) phenotype. The L1 larvae were paralyzed and kinky as soon as they hatched; adults moved slightly better but were extreme forward and backward coilers (Table I). In addition, *f131* animals were small, thin, transparent, and slow-growing and had a brood size half that of the wild type. From the interactions with the *nDF17* deletion, fully uncovering the region, that lead to a stronger phenotype, *i.e.* a high frequency of lethals (Table I), and the molecular evidence presented below, we believe that these three uncoordinated alleles, *e189, f120,* and *f131,* are partial loss-of-function alleles.

Two unc-32 Alleles Are Lethal, and the Locus Is Complex—The two further alleles, *f121* and *f123,* isolated in the screen are zygotic lethal. *f123* is equivalent to the deficiency in all interallelic combinations and therefore behaves as a null (Table I). Additionally, it is suppressible by the amber nonsense suppressor *sup-7* (14), suggesting the presence of a premature stop codon. *f123* homozygotes arrest as late embryos, often with vacuolization of the intestinal region. This null allele is associated with a strict maternal effect lethality, giving rise to severe cell adhesion defects in the early embryos. In addition, we cannot exclude a sterility owing to lack of production of mature gametes. The weaker lethal allele *f121* arrests at the larval L2, L3, or L4 stage, and before their premature death the movement of the larvae appears close to that of wild type. In contrast to *f123*, *f121* almost fully complements the mild Unc alleles *e189* and *f120* and is fully viable when combined with *e189, f120,* and *f131. f121*/*f131* larvae move better than *f131* homozygotes but are still highly impaired in their locomotion. Adults, however, are only backward ventral coilers, just like *e189.* Hence, the lethality of *f121*, unlike that of *f123,* is rescued by each of the three Unc alleles, and conversely the Unc phenotype of these three alleles is partly rescued by *f121.* Such interallelic complementation defines *unc-32* as a complex locus, a property often associated with alternative splicing.

unc-32 Encodes a 100-kDa a Subunit of V-ATPase—Genetic mapping had previously placed *unc-32* between *sma-2* and *lin-12* in the center of *LG III* (13). Transformation rescue of *unc-32(e189)* by microinjection of cosmids from the candidate region of the physical map showed that the gene is contained within both ZK637 and T05B2[2] (Fig. 2*A*). We further refined the position of the gene by transformation rescue to a region of 9 kilobases (E5-BK, E5-NS, and E5-lin-9 constructs; see Fig. 2*B*), corresponding to the predicted gene ZK637.8. This gene encodes a protein similar in sequence to V-ATPase 100 kDa *a* subunits. Constructions in which part of ZK637.8 was deleted or in which a premature stop codon was introduced into the coding sequence of the gene failed to rescue the uncoordinated phenotype (E5-ΔPac and E5-BK* constructs; see Fig. 2*B*). Although a rescue of the zygotic lethality of *f123* and *f121* was obtained, the transformants were sterile (Table I and Fig. 2), possibly because of an absence of germ line expression of the transgenes as previously reported (23). Occasionally, the transforming plasmids rescued the various phenotypes, Unc, lethal, maternal effect lethal, and defect in gametogenesis, in a complex pattern. A particular transgenic array could rescue the Unc but not the lethal phenotype, whereas another, containing the same constructs, would do the opposite. This again presumably reflects the complexity of the *unc-32* locus (see below).

Sequencing of the ZK637.8 genomic region from four mutant alleles revealed different molecular alterations: *e189* and *f120* contain a point mutation in the splice acceptor site of exon 4b, *f131* in the splice donor site of exon 6 (see below), and *f121* is a G to A transition in exon 6 (Table I). Together, these results confirm the identity of *unc-32* as ZK637.8. The gene is the first in an operon containing two others, ZK637.9 and ZK637.10. The former encodes a protein similar in sequence to thiamine

[2] J. Sulston, personal communication.

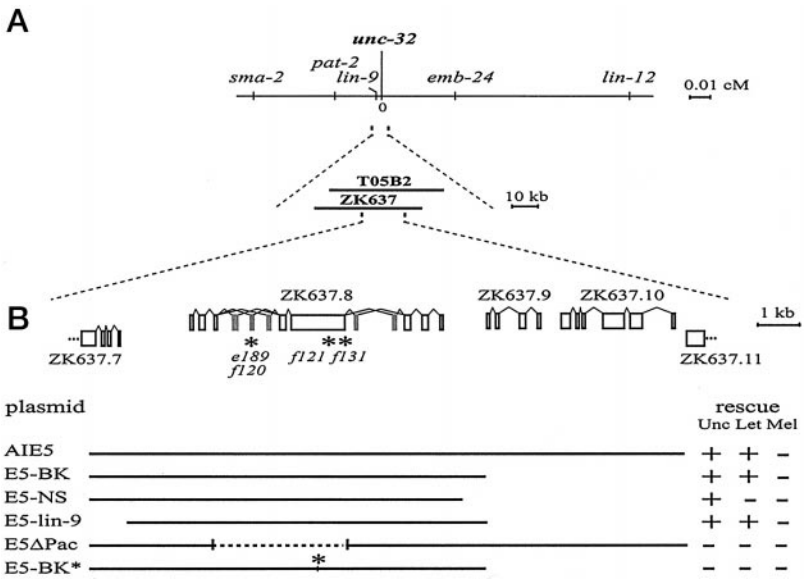

FIG. 2. **Characterization of the *unc-32* gene.** *A*, Genetic and physical maps of the *unc-32* region. The cosmids ZK637 and T05B2 were shown by J. Sulston to rescue the uncoordinated phenotype of *unc-32(e189)* mutants. *B*, *top*, predicted and corrected genes in the indicated part of cosmid ZK637, with invariant exons indicated by *open boxes*; alternative exons in ZK637.8 are *gray*. The direction of transcription of ZK637.7, which corresponds to *lin-9* (35), and of ZK637.11 is from *right* to *left*. For the other genes, it is from *left* to *right*. The positions of the point mutations of three *unc-32* alleles are shown with *asterisks*: *e189* is a G-A transition in the acceptor splice site in exon 4b; *f131* is a G-A transition in the donor splice site in exon 6; and *f121* is a G-A transition in exon 6. *Bottom*, inserts of the different clones tested for their ability to rescue the different *unc-32* phenotypes. *Let*, lethal; *Mel*, maternal effect lethal. Transformation rescue experiments were performed in either *e189*, *f131* or the lethal alleles, and only a summary of the rescue is presented. For example, the *AIE5* construct rescued the movement and fertility defects of *f131* mutants but was associated with the appearance of an *f123*-like maternal effect lethality such that 85% of the eggs were dead. The clone *E5ΔPac* contains a deletion corresponding to the major part of ZK637.8. The clone *E5-BK** harbors a point mutation resulting in a frame shift in the sixth exon and consequently a premature stop codon, the position of which is marked by an *asterisk*.

pyrophosphokinase (24). The latter encodes a putative mitochondrial thioredoxin reductase. By RT-PCR, we showed that these three genes are expressed, and that although the transcript of ZK637.8 is trans-spliced to SL1, both ZK637.9 and ZK637.10 can be trans-spliced to SL2 (data not shown; Ref. 25), as is normally the case for genes in *C. elegans* operons (26). Although in many operons the genes are functionally related, no evident link is apparent so far between the three genes of the *unc-32* operon.

unc-32 Transcripts Are Alternatively Spliced—Sequencing of existing cDNAs and specific RT-PCR products revealed a number of alternative splice patterns for the *unc-32* transcripts (Fig. 3, *A* and *B*). All *unc-32* transcripts analyzed contain 10 exons. There are at least six different transcripts, using three variants for exon 4 (a–c) and two for exon 7 (a and b). Although the alternative exons 4a–c do not encode peptides that closely resemble sequences found in other V-ATPases, they occur at a position in the sequence where an alternative splicing pattern has also been observed for the murine *a1* gene (9), suggesting that they may have an important functional role. Significantly, the mutation in *unc-32(e189)* is a G to A transition in the consensus splice acceptor site of exon 4b, meaning that in this allele, only the alternative transcripts with exon 4b should be affected. Consistent with this, anomalous *unc-32* transcripts were found by RT-PCR amplification of the fourth exon region of total RNA extracted from *unc-32(e189)* homozygotes (Fig. 3*C*). Cloning and sequencing of the forms that contain exon4b in *e189* homozygotes revealed the addition of 22 bp upstream of the exon 4b because of the use of a new acceptor splicing site. The resulting anomalous transcript contains a premature stop codon and would encode a nonfunctional protein. The other alternative forms containing exons 4a and 4c are still produced (Fig. 3*C*).

The alternative seventh exons, 7a and 7b, are similar in sequence and presumably arose from a recent duplication, because a single exon of similar sequence is present in other species (Fig. 3A). The *f131* allele contains a G to A transition at the consensus splice donor site of the sixth exon, which affects the splicing of these alternative seventh exons (Fig. 3*C*). The large anomalous *unc-32* transcript found by RT-PCR amplification of the seventh exon region of total RNA extracted from *unc-32(f131)* homozygotes was sequenced. This revealed the presence of intron 6 between exon 6 and exon 7a. Because a stop codon is present in the sixth intron, the resulting mutant protein would be truncated in the eighth transmembrane domain. The other band with a size similar to that of wild type (Fig. 3*C*) corresponds to a mixture of alternate transcripts using cryptic donor sites. Because this allele is a hypomorph, not a null, we expect that some of these transcripts can be translated into a near full-length product that would differ from the wild type only in the relatively poorly conserved region around the exon 6-exon 7 junction.

V-ATPase a Subunits Are a Large and Highly Conserved Family—UNC-32 belongs to a family of highly conserved V-ATPase *a* subunit proteins. In other species there are multiple genes encoding V-ATPases *a* subunits. There are three genes in mice (9) and other vertebrates, and searches in the available *Drosophila* genomic sequence indicate that there are at least four members in the fly.[3] In *C. elegans*, in addition to UNC-32, there are three other members of the family, named VHA-5, VHA-6, and VHA-7. Each of the four genes has a distinctive intron-exon structure. We could not detect alternative splicing of these genes, as judged by the analysis of several cDNAs.[4] The existence of shared intron-exon boundaries among the four nematode genes and the two human genes for which data are available supports their common ancestry (Fig. 4*A*). Each of these putative proteins contain nine potential membrane-spanning regions (Ref. 27 and Fig. 4*B*). The sequence identity (and

[3] GenBank™ accession numbers AAD34751, AAD34771, AAF53116, and AAF55550.

[4] D. Thierry-Mieg and Y. Kohara, unpublished results.

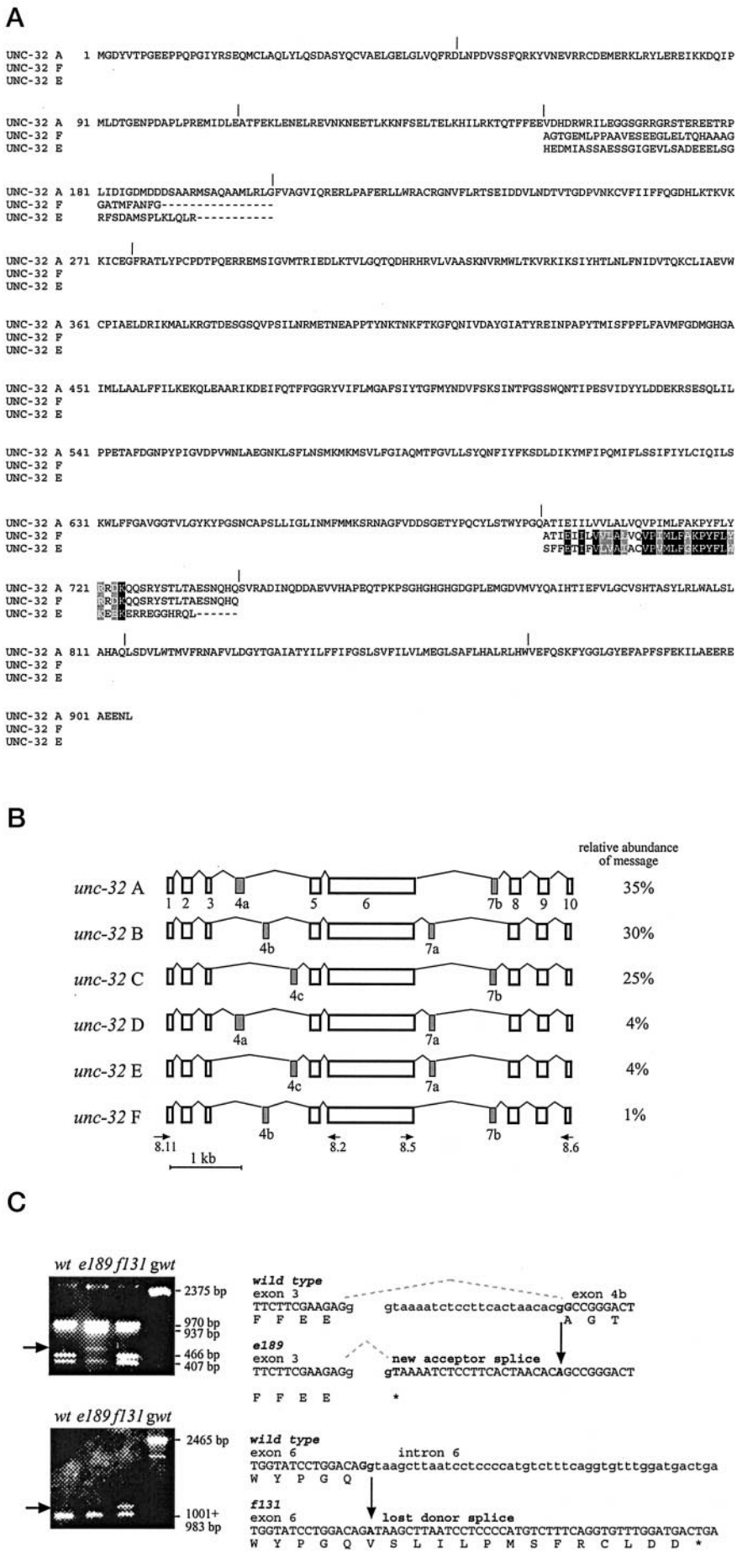

FIG. 3. **The *unc-32* gene encodes several isoforms of V-ATPase *a* subunit.** The different *unc-32* transcripts were determined by sequencing different cDNA clones and RT-PCR products, and the respective proteins are named as follows: *UNC-32 A–F* contain, respectively, exons 4a/7b, 4b/7a, 4c/7b, 4a/7a, 4c/7a, and 4b/7b. Their GenBank accession numbers are, respectively, AF320899, AF320900, AF320901, AF320902, AF320903, and AF320904. Their predicted molecular masses, in kilodaltons, are, respectively, 103.4, 100.7, 102, 102.8, 101.3, and 101.1. Note that the different UNC-32 proteins differ from both predicted ZK637.8a and 8b proteins. *A*, UNC-32 A, F, and E protein sequences. Apart from the alternative exons (of variable size, hence marked with *dashes*), the sequences are identical and are not shown for the sake of clarity. Although the fourth exons are not similar, exons 7b and 7a in the variants F and E, respectively, are highly similar. Residues that are identical are *black*; those that are similar are *gray*. The *vertical lines* above the sequences mark the corresponding exon-intron boundaries. *B*, exon-intron structure of the six different alternative splice variants. The primers used for the RT-PCR presented in *C* are also shown. Alternative exons are *gray*. The relative abundances of different transcripts were evaluated after digestion of RT-PCR products made from mRNA from mixed stage wild-type worms and are shown as *percentages*. *C*, different patterns of mRNA splicing are revealed by RT-PCR in the *unc-32* mutants, and their sequences are presented. *Top panel*, analysis of the *e189* mutation. On the *left*, the amplicon containing the fourth exon was amplified with the primers 8.11 and 8.2 and digested with *Hae*III. Transcripts containing exon 4a (970 bp) or 4c (937 bp) are not cut, whereas transcripts containing exon 4b are cut in three fragments (466, 407, and 49 bp; the 49-bp band is not visible in this image). The *e189* mutation removes a *Hae*III site, which is normally formed by the splicing of exon 3 to 4b, and results in the appearance of a larger band (*arrow*) and the loss of the 49-bp band (not shown). The *band* at *2375 bp* results from direct PCR amplification of genomic DNA, shown in *lane gwt*. On the *right* are presented the sequences of both wild-type and *e189* transcripts. The *e189* mutation removes the exon 4b acceptor splicing site; a new acceptor splicing site, located 22 bp upstream of the normal one, is used, introducing a stop codon. *Bottom panel*, analysis of the *f131* mutation. On the *left*, the amplicon including the seventh exon was amplified with primers 8.5 and 8.6. Transcripts containing exon 7a (983 bp) or 7b (1001 bp) are not resolved. The *f131* mutation in the splice donor site of exon 6 provokes the apparition of a larger band (*arrow*). The *band* at *2465 bp* results from direct PCR amplification of genomic DNA, shown in *lane gwt*. On the *right* are presented the sequences of both wild-type and *f131* transcripts. The *f131* mutation removes the exon 6 donor splicing site; therefore, intron 6 is not spliced out and fuses to exon 7a, resulting in the introduction of stop codons. The *bottom band* is a mixture of various transcripts using cryptic splice sites.

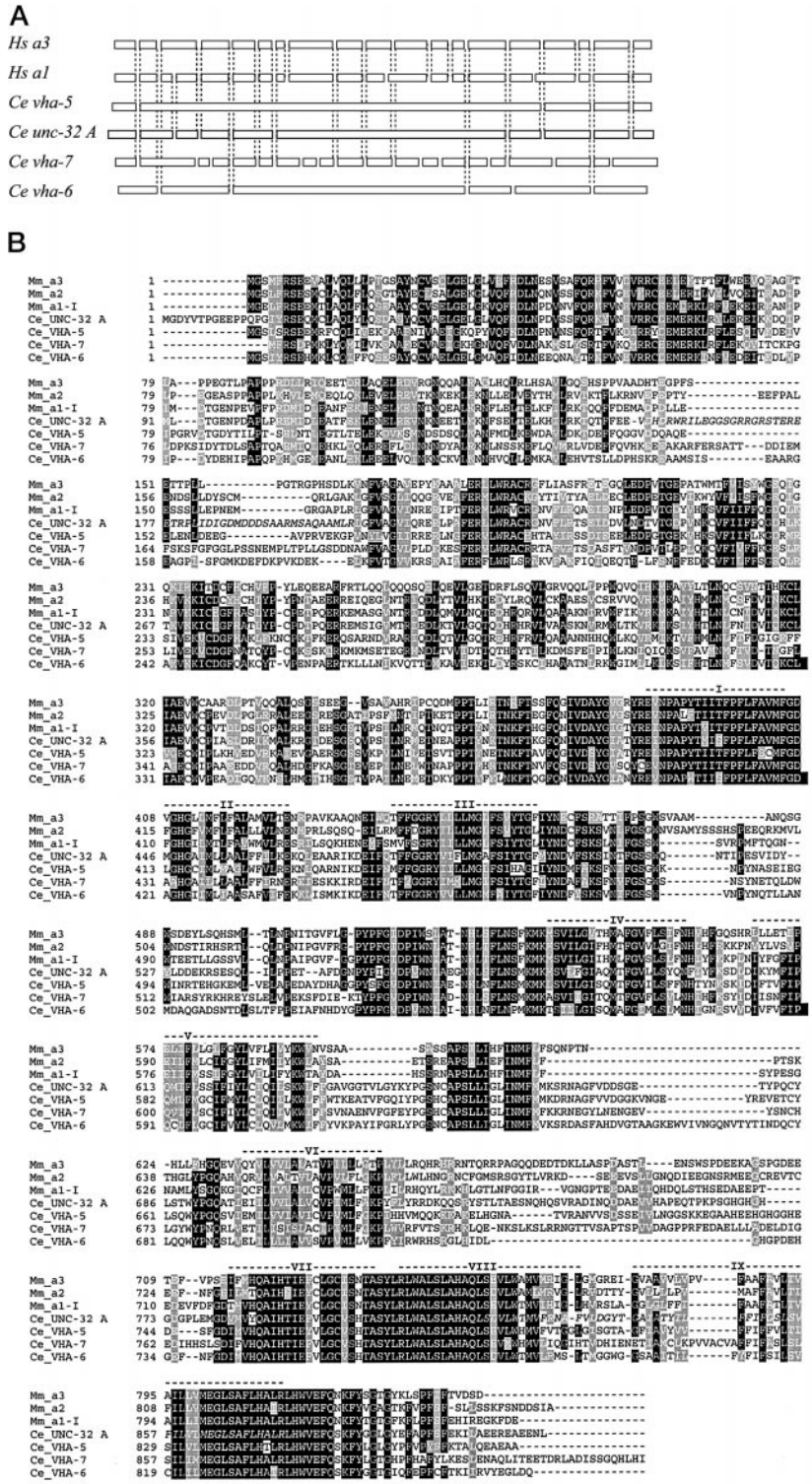

FIG. 4. **Structure and sequence alignment of the four *C. elegans* V-ATPase *a* subunit genes.** *A*, Intron-exon structure of the four *C. elegans* V-ATPase *a* subunit genes and the *a1* and *a3* human isoforms. The *open boxes* represent exons drawn to scale, with the exception of the largest ones. Note that the distance between the exons is not to scale. The *dotted lines* show intron-exon boundaries shared between two or more genes. *B*, amino acid sequence alignment of the *a* subunit proteins from mouse and nematode. Alignment of the three mouse *a* subunit protein sequences (*a1–a3*; Ref. 9) and the four nematode sequences (the UNC-32 A splice variant described here, VHA-5 (AAA81682), VHA-6 (CAA90758), and VHA-7, is shown. The latter is an edited version of the Genefinder prediction (see "Experimental Procedures"). The position of the nine putative transmembrane helices is indicated with *dotted lines*, and the amino acids encoded by alternative exons 4a and 7b are in *italics*.

similarity) between UNC-32A and other nematode and murine *a* subunit proteins are as follows: VHA-5, 50% (67%); VHA-6, 50% (65%); VHA-7, 46% (62%); a1, 55% (69%); a2, 43% (60%); and a3, 37% (55%). Phylogenetic analysis of the different members of this family indicates, in agreement with previous suggestions based on a more limited analysis (9), that the multiple genes appeared independently in vertebrates and invertebrates subsequent to their divergence from a common ancestor (Fig. 5). Interestingly, although the UNC-32 protein is closer in sequence to the vertebrate *a1* protein than to the other nematode proteins, the *vha-7* gene shares more intron-exon boundaries with the human *a1* and *a3* genes than other nematode genes.

V-ATPase a Subunits Have Specific Expression Domains—The expression pattern of each of the four V-ATPase *a* genes in *C. elegans* was analyzed by *in situ* hybridization. Two of the genes, *vha-6* and *vha-7,* have very limited and nonoverlapping domains of expression, whereas the two others, *unc-32* and *vha-5,* are expressed more broadly, yet each with a distinctive pattern (Fig. 6). Transcripts for *unc-32* were detected at high levels in early embryos (Fig. 6*A*) from the one-cell stage to the end of gastrulation. In larvae and adults, transcripts were

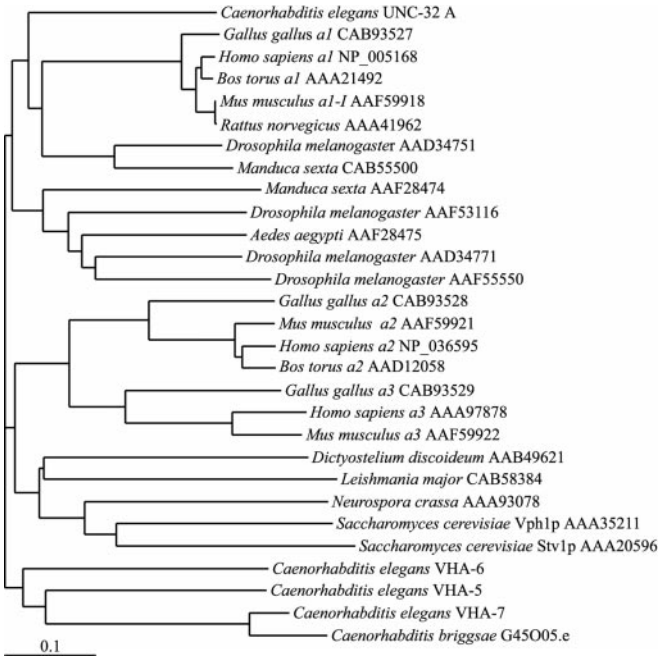

FIG. 5. **Evolutionary relationship between the V-ATPase 100-kDa *a* subunit proteins in different eukaryotes.** The phylogenetic tree was constructed using the *C. elegans* sequences shown in Fig. 4*B*, as well as those of other organisms present in publicly accessible data bases. The structure of the tree is compatible with an independent appearance of different proteins subsequent to the evolutionary divergence of vertebrates and invertebrates and reflects the fact that the UNC-32 protein is more closely related to the vertebrate *a1* protein than to other nematode proteins. *Scale bar*, 0.1 nucleotide substitutions per site.

clearly present in the gonad (Fig. 6, *C* and *D*), in the intestine, in many neurons in the head, and in motoneurons in the ventral cord (Fig. 6*D*). This expression pattern is consistent with the expression of a LacZ reporter gene driven by the upstream region of the *unc-32* gene (28).[5] Similarly, *vha-5* showed precocious and widespread embryonic expression (Fig. 6*E*), followed by a more restricted pattern in larvae and adults. mRNAs are present in many tissues where *unc-32* is also expressed, but the detailed patterns were different, because the gene is also expressed in the excretory cell and canals (Fig. 6*F*) and probably some somatic gonad cells (Fig. 6, *G* and *H*). In contrast, expression of *vha-6* started late, in the intestine of the 2-fold embryo (Fig. 6*I*). It was strong but limited to the intestinal cells at all stages (Fig. 6, *J–L*). The *vha-7* gene was silent during embryogenesis and the first larval stages (Fig. 6*M*). Its expression was first observed at the L3 stage and was limited to areas of the developing and mature gonad corresponding to the spermatheca (Fig. 6*N*) and the region where meiotic oocytes are in the pachytene and early diakinesis stages. It vanished abruptly in mature oocytes (Fig. 6, *O* and *P*). Although the patterns of expression of these other genes are suggestive, their functions remain to be investigated.

unc-32 Transcripts Containing Exon 4b Are Specifically Expressed in the Nervous System—In situ hybridization with the *unc-32* gene revealed a large domain of expression, especially in the embryo where the expression is ubiquitous. To test the hypothesis that *unc-32* alternative transcripts have tissue-specific expression patterns, we made a fusion construct, *unc-32*e4::GFP, in which GFP was inserted into the alternative exon 4b. GFP expression in wild-type strains carrying this construct was restricted to the nervous system (Fig. 6*Q*), including all the motoneurons in the ventral cord, the dorsal, the lateral, and all the sublateral cords, PLML/R, PVM, ALML/R, AVM, ALNL/R, PLNL/R, PVDL/R, PDEL/R, SDQL/R, and almost all the neurons in the head.

DISCUSSION

unc-32(e189) Is Affected in the Motor Circuit—The analysis of the uncoordinated phenotype of *unc-32(e189)* suggested a possible defect in specific classes of motoneurons (29). Serial reconstruction of the anterior part of the ventral nerve cord behind the retrovesicular ganglion in the mutant, however, failed to show any wiring defect in the motoneurons. They revealed instead that synaptic vesicles had an abnormal and distinctive morphology compared with wild type.[6] This strongly suggests that the function rather than the structure of the motor circuit is compromised in *unc-32(e189)* mutants. Consistent with this, mutant animals are resistant, at a low but consistent level, to inhibitors of acetylcholine esterase, as are mutants in other genes involved in acetylcholine synthesis, packaging, release, or signal reception (30). We have unambiguously identified *unc-32* as being a nematode V-ATPase *a* subunit and show that *unc-32* is expressed in some motoneurons and is thus likely to be directly responsible for the uncoordinated phenotype. We propose that in the *unc-32(e189)* mutant, V-ATPases are not functional in motoneurons, leading to a lack of import of the acetylcholine neurotransmitter into synaptic vesicles, a function that as been previously suggested for V-ATPase (4, 5).

Alternative Splice Forms Are Expressed in Motoneurons—There are at least six alternative transcripts for the *unc-32* gene. The *unc-32(e189)* mutation affects the generation of two of the six transcripts, namely those containing exon 4b. In other terms, in the *e189* allele, there is no transcript with exon 4b, or if a transcript contains exon 4b, it is nonfunctional. One hypothesis would be that the isoforms with exon 4b are required for synaptic vesicle function in motoneurons but are dispensable in other tissues. Alternatively, motoneurons could force the splicing apparatus to always include exon 4b, thus generating a null allele in this cell type. The fact that a GFP fusion in exon 4b is specifically expressed in the nervous system favors the latter hypothesis.

Other Alternative Splice Forms Are Required during Development—The *unc-32(f121)* and the null *unc-32(f123)* alleles are lethals, so certain UNC-32 isoforms must be absolutely required for viability. Because *unc-32(f121)* corresponds to a point mutation in an exon common to all isoforms and will affect all alternative *unc-32* transcripts, it remains to be determined which UNC-32 isoforms have an essential function. The lethality of those alleles is, however, consistent with the fact that *unc-32* is widely expressed in early embryogenesis. Rescue of the *e189* uncoordinated phenotype by the *f121* lethal allele suggests that at least one of the *f121* mutant forms is partially functional in the synaptic vesicles of the motor circuit. The *f121* mutation results in a Glu for Gly substitution in the third transmembrane domain at a position that is not fully conserved between species. In motoneurons, this substitution in transcripts containing exon 4b might not perturb the function of the V-ATPase.

Our results have revealed that the function and regulation of *unc-32* are complex. First, there are at least six alternative transcripts for this gene. Second, the three uncoordinated alleles are affected in the generation of these alternative transcripts. Third, the different mutant alleles display distinct phenotypes and unexpected interallelic complementation. These could be due to perturbations of an equilibrium between

[5] A. Lynch and I. Hope, personal communication.

[6] J. White and E. Southgate, personal communication.

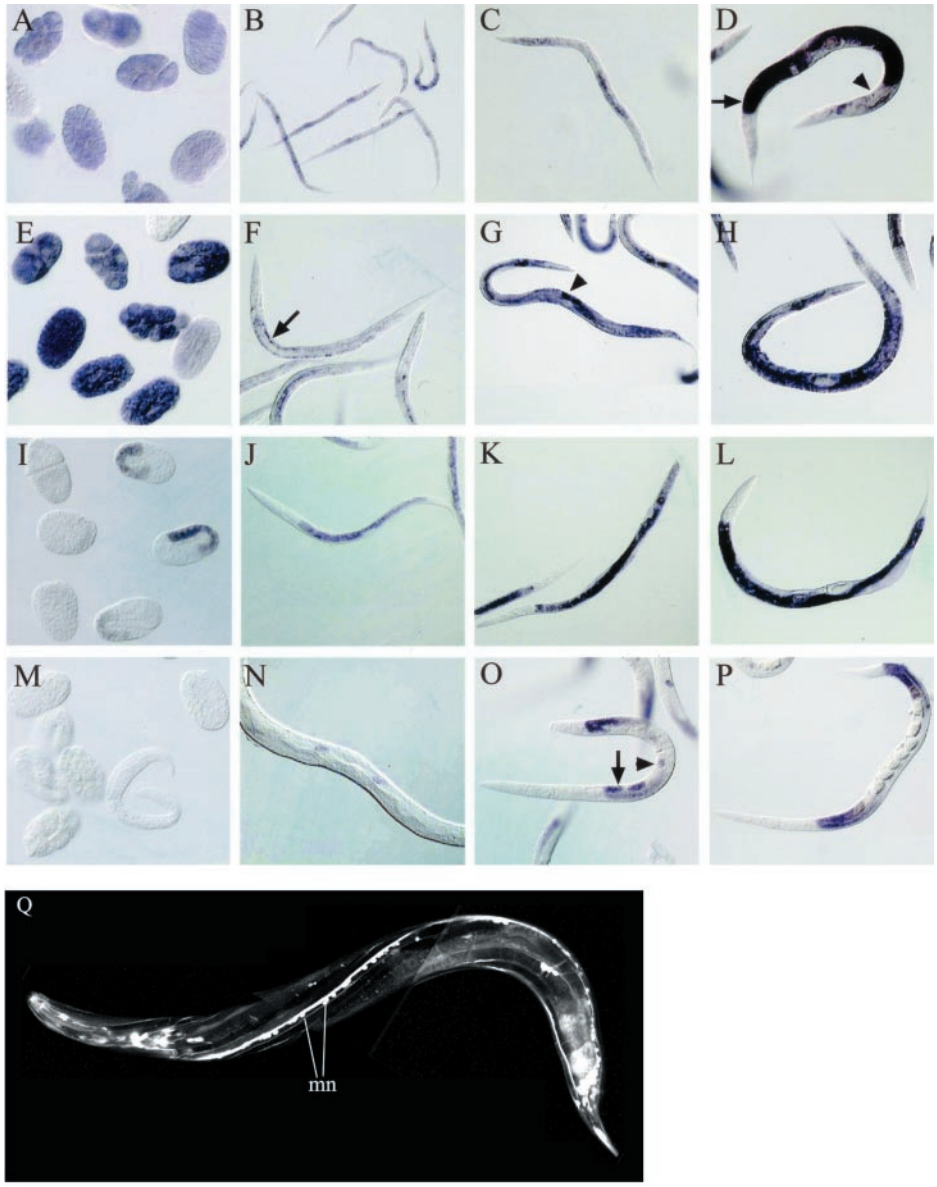

FIG. 6. **Expression pattern of the four V-ATPase *a* subunit genes in the nematode.** *In situ* hybridization images on embryonic, larval, and adult worms with *unc-32* (*A–D*), *vha-5* (*E–H*), *vha-6* (*I–L*), and *vha-7* (*M–P*) are shown. *unc-32* is expressed early and widely in the embryo (*A*); expression is seen in the intestine (not shown), the gonad (*arrow*), and some neurons in the head as well as motoneurons (*arrowhead*) in the ventral cord in larvae (*B* and *C*) and adult (*D*). *vha-5* is also expressed early and widely in the embryo (*E*); expression is seen in the excretory canal (*arrow*) and some somatic gonad cells (*arrowhead*) in larvae (*F* and *G*) and adult (*H*). *vha-6* is expressed only in the intestine, from the embryonic coma stage (*I*) through larvae (*J* and *K*) and adult (*L*). *vha-7* is not expressed in embryos and young larvae (*M*). Expression starts at the L3-L4 larval stage (*N*) and is seen in early oocytes (*arrow*) and in the spermateca (*arrowhead*) in L4 (*O*) and adult (*P*). *Q*, Confocal images of an adult transgenic worm, carrying the construct *unc-32*e4::GFP. Expression is limited to the nervous system. *mn*, motoneurons.

alternatively spliced forms of *unc-32*, each of which would encode a protein with a specific function. The *smg-2* suppressor is involved in mRNA surveillance (15) and has been shown to modulate the level of alternatively spliced mRNAs (16). Significantly, the lethality of *f121* was rescued by *smg-2,* but *unc-32(f121); smg-2(e2008)* adults did not lay eggs (data not shown).

Four V-ATPase a Subunits in the Nematode—The *unc-32* gene corresponds to one of four nematode V-ATPase *a* subunits. Just like in vertebrates, where different *a* subunit isoforms are targeted to different tissues (9, 31), each of the *C. elegans* V-ATPase *a* genes has a specific, largely nonoverlapping, domain of expression. Tissue-specific expression of alternatively spliced forms of the murine *a1* gene has also been reported (9). Furthermore, in yeast, the two *a* subunit isoforms have been implicated in the differential intracellular localization of the V-ATPase complex (8). This suggests that different isoforms of the *a* subunit, arising from alternative splicing or the expression of different genes, or both, have specific patterns of expression, both in terms of tissue distribution and targeting to individual intracellular membrane compartments. In the nematode, further variation of the structure of the V-ATPase is possible, because three *c* subunit genes (*vha-1*, *-2*, and *-3*) have been described (32–34). Given their conservation through evolution, future studies using *C. elegans* may allow a deeper understanding of the enormous potential for variation in the structure, function, and regulation of V-ATPases and their association with different pathologies in vertebrates (11).

Acknowledgments—We thank C. Couillault, L. Guiraud and F. Jamen for their participation in this project, M. Craxton, A. Fire, and J. Kimble for the generous gift of clones and strains, I. Hope, A. Lynch, J. Sulston, and J. White for sharing unpublished results, C. Goridis and J.-F. Brunet for critical reading of the manuscript, and J. Thierry-Mieg, M. Weill, and J. Demaille for their support.

REFERENCES

1. Nelson, N. (1992) *J. Exp. Biol.* **172,** 149–153
2. Forgac, M. (1999) *J. Biol. Chem.* **274,** 12951–12954
3. Futai, M., Oka, T., Sun-Wada, G., Moriyama, Y., Kanazawa, H., and Wada, Y. (2000) *J. Exp. Biol.* **203,** 107–116
4. Varoqui, H., and Erickson, J. D. (1996) *J. Biol. Chem.* **271,** 27229–27232
5. Gasnier, B. (2000) *Biochimie (Paris)* **82,** 327–337
6. Nelson, N., Perzov, N., Cohen, A., Hagai, K., Padler, V., and Nelson, H. (2000) *J. Exp. Biol.* **203,** 89–95
7. Forgac, M. (2000) *J. Exp. Biol.* **203,** 71–80
8. Manolson, M. F., Wu, B., Proteau, D., Taillon, B. E., Roberts, B. T., Hoyt, M. A., and Jones, E. W. (1994) *J. Biol. Chem.* **269,** 14064–14074
9. Nishi, T., and Forgac, M. (2000) *J. Biol. Chem.* **275,** 6824–6830
10. Li, Y. P., Chen, W., Liang, Y., Li, E., and Stashenko, P. (1999) *Nat. Genet.* **23,** 447–451
11. Frattini, A., Orchard, P. J., Sobacchi, C., Giliani, S., Abinun, M., Mattsson, J. P., Keeling, D. J., Andersson, A. K., Wallbrandt, P., Zecca, L.,

Notarangelo, L. D., Vezzoni, P., and Villa, A. (2000) *Nat. Genet.* **25,** 343–346
12. Kornak, U., Schulz, A., Friedrich, W., Uhlhaas, S., Kremens, B., Voit, T., Hasan, C., Bode, U., Jentsch, T. J., and Kubisch, C. (2000) *Hum. Mol. Genet.* **9,** 2059–2063
13. Brenner, S. (1974) *Genetics* **77,** 71–94
14. Wills, N., Gesteland, R. F., Karn, J., Barnett, L., Bolten, S., and Waterston, R. H. (1983) *Cell* **33,** 575–583
15. Pulak, R., and Anderson, P. (1993) *Genes Dev.* **7,** 1885–1897
16. Morrison, M., Harris, K. S., and Roth, M. B. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 9782–9785
17. Mello, C. C., Kramer, J. M., Stinchcomb, D., and Ambros, V. (1991) *EMBO J.* **10,** 3959–3970
18. Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159
19. Krause, M., and Hirsh, D. (1987) *Cell* **49,** 753–761
20. Huang, X. Y., and Hirsh, D. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 8640–8644
21. Tabara, H., Motohashi, T., and Kohara, Y. (1996) *Nucleic Acids Res.* **24,** 2119–2124
22. Heinemann, T., Bulwin, G. C., Randall, J., Schnieders, B., Sandhoff, K., Volk, H. D., Milford, E., Gullans, S. R., and Utku, N. (1999) *Genomics* **57,** 398–406
23. Mello, C., and Fire, A. (1995) *Methods Cell Biol.* **48,** 451–482
24. Nosaka, K., Onozuka, M., Nishino, H., Nishimura, H., Kawasaki, Y., and Ueyama, H. (1999) *J. Biol. Chem.* **274,** 34129–34133
25. Zorio, D. A., Cheng, N. N., Blumenthal, T., and Spieth, J. (1994) *Nature* **372,** 270–272
26. Spieth, J., Brooke, G., Kuersten, S., Lea, K., and Blumenthal, T. (1993) *Cell* **73,** 521–532
27. Leng, X. H., Nishi, T., and Forgac, M. (1999) *J. Biol. Chem.* **274,** 14655–14661
28. Lynch, A. S., Briggs, D., and Hope, I. A. (1995) *Nat. Genet.* **11,** 309–313
29. Chalfie, M., and White, J. (1988) in *The Nematode Caenorhabiditis Elegans* (Wood, W. B., ed) pp. 337–391, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
30. Nguyen, M., Alfonso, A., Johnson, C. D., and Rand, J. B. (1995) *Genetics* **140,** 527–535
31. Peng, S. B., Crider, B. P., Xie, X. S., and Stone, D. K. (1994) *J. Biol. Chem.* **269,** 17262–17266
32. Oka, T., Yamamoto, R., and Futai, M. (1997) *J. Biol. Chem.* **272,** 24387–24392
33. Oka, T., Yamamoto, R., and Futai, M. (1998) *J. Biol. Chem.* **273,** 22570–22576
34. Oka, T., and Futai, M. (2000) *J. Biol. Chem.* **275,** 29556–29561
35. Ferguson, E. L., and Horvitz, H. R. (1989) *Genetics* **123,** 109–121